



ENTERED

PCT09

RAW SEQUENCE LISTING

DATE: 03/05/2002 ^{PS}

PATENT APPLICATION: US/09/831,142B

TIME: 14:22:53

Input Set : A:\Wcm-69us.app

Output Set: N:\CRF3\03052002\I831142B.raw

```

3 <110> APPLICANT: CAMPBELL, ANTHONY KEITH
5 <120> TITLE OF INVENTION: PROTEIN AND DNA CODING THEREFOR
7 <130> FILE REFERENCE: WCM.69.US
9 <140> CURRENT APPLICATION NUMBER: 09/831,142B
10 <141> CURRENT FILING DATE: 2001-05-07
12 <150> PRIOR APPLICATION NUMBER: PCT/GB99/03654
13 <151> PRIOR FILING DATE: 1999-11-05
15 <150> PRIOR APPLICATION NUMBER: GB 9824357.9
16 <151> PRIOR FILING DATE: 1998-11-07
18 <160> NUMBER OF SEQ ID NOS: 42
20 <170> SOFTWARE: PatentIn Ver. 2.1
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 870
24 <212> TYPE: DNA
25 <213> ORGANISM: Pholas dactylus
27 <220> FEATURE:
28 <221> NAME/KEY: CDS
29 <222> LOCATION: (30)..(704)
31 <400> SEQUENCE: 1
32 gaattcggca cgagtcggaa aagaacaaa atg gct tgt atc gtt ttc gtt gct 53
33                               Met Ala Cys Ile Val Phe Val Ala
34                               1           5
36 ctt gtc gct cta tgc tta atg caa ccg ggt tcc ggt gag gaa gta caa 101
37 Leu Val Ala Leu Cys Leu Met Gln Pro Gly Ser Gly Glu Glu Val Gln
38      10           15           20
40 tgc gcg atg aat tgg aca caa gct aat gaa tat gtg ttc aac gtg gac 149
41 Cys Ala Met Asn Trp Thr Gln Ala Asn Glu Tyr Val Phe Asn Val Asp
42  25           30           35           40
44 tgg atg acc att ttc atc tac gac tat ggc gct caa gag caa ctg tac 197
45 Trp Met Thr Ile Phe Ile Tyr Asp Tyr Gly Ala Gln Glu Gln Leu Tyr
46           45           50           55
48 gaa gat cgg gct ttg ggg ctg tgt cgg att gaa cgg gcc ggc cca ggt 245
49 Glu Asp Arg Ala Leu Gly Leu Cys Arg Ile Glu Arg Ala Gly Pro Gly
50      60           65           70
52 acc aca aaa gcc gtc tgg att aac tgg agt aac gac acg cag tca tgt 293
53 Thr Thr Lys Ala Val Trp Ile Asn Trp Ser Asn Asp Thr Gln Ser Cys
54      75           80           85
56 gta aca aga aaa aca atc ttc ttc gag gtt ggt gga gaa att gcc cgg 341
57 Val Thr Arg Lys Thr Ile Phe Phe Glu Val Gly Gly Glu Ile Ala Arg
58      90           95          100
60 cta gtt gac tac aga cca cag gaa gac gga act gag aaa act ttt aca 389
61 Leu Val Asp Tyr Arg Pro Gln Glu Asp Gly Thr Glu Lys Thr Phe Thr
62 105           110           115           120

```

RAW SEQUENCE LISTING

DATE: 03/05/2002

PATENT APPLICATION: US/09/831,142B

TIME: 14:22:53

Input Set : A:\Wcm-69us.app

Output Set: N:\CRF3\03052002\I831142B.raw

```

64 aga aaa ttc tct agc aaa atg cca ggc act tac atg ctt atg gac gtg 437
65 Arg Lys Phe Ser Ser Lys Met Pro Gly Thr Tyr Met Leu Met Asp Val
66          125          130          135
68 tgc gct aca agg gac gct gat gat aaa tgc atc gaa ggc aca att gtg 485
69 Cys Ala Thr Arg Asp Ala Asp Asp Lys Cys Ile Glu Gly Thr Ile Val
70          140          145          150
72 gtg aca gtc agg gtg tcc cta tat gac gaa gat aac aat ggt gta atg 533
73 Val Thr Val Arg Val Ser Leu Tyr Asp Glu Asp Asn Asn Gly Val Met
74          155          160          165
76 gat gaa ggt aag gtg att cca tct gag aca atc gag gat gat atc aag 581
77 Asp Glu Gly Lys Val Ile Pro Ser Glu Thr Ile Glu Asp Asp Ile Lys
78          170          175          180
80 gac tgt ggg ctc tta gac caa gat gtt gaa ctc gat tat acg tgg act 629
81 Asp Cys Gly Leu Leu Asp Gln Asp Val Glu Leu Asp Tyr Thr Trp Thr
82 185          190          195          200
84 caa aac gag tgt gat cta cca gac aca gta gac gag gct gaa gac aca 677
85 Gln Asn Glu Cys Asp Leu Pro Asp Thr Val Asp Glu Ala Glu Asp Thr
86          205          210          215
88 ccg tca gaa act gga gaa ttc ttc tgg tagatctatc agactacttt 724
89 Pro Ser Glu Thr Gly Glu Phe Phe Trp
90          220          225
92 tatcagcagg acaactggtc gttaccagac acctataacg tgctctcctc aataatgtgt 784
94 aaaacagaaa taatcgatag aatattgaaa ataaaatggt aataaacact ggttgaaata 844
96 tgaaaaaaaaa aaaaaaaaaa ctcgag 870
98 <210> SEQ ID NO: 2
99 <211> LENGTH: 816
100 <212> TYPE: DNA
101 <213> ORGANISM: Pholas dactylus
103 <400> SEQUENCE: 2
104 gaattcggca cgaggga aaa gaacaaaatg gcttgtatcg ttttcgttgc tcttgctcgt 60
105 ctatgcttaa tgcaaccggg ttccgggtgag gaagtacaat gcgcgatgaa ttggacacaa 120
106 gctaataaat atgtgttcaa cgtggactgg atgaccattt tcatctacga ctatggcgct 180
107 caagagcaac tgtacgagga tcgggctttg gggctgtgtc ggattgaacg ggccggccca 240
108 ggtaccacaa aagccgtctg gattaactgg agtaacgaca cgcagtcgat tgtaacaaga 300
109 aaaacaatct tcttcgaggt tgggtggagaa attgccggc tagttgacta cagaccacag 360
110 gaagacggaa ctgagaaaac ttttacaaga aaattctcta gcaaaatgcc aggcacttac 420
111 atgcttatgg acgtgtgcgc tacaaggagc gctgatgata aatgcacga aggcacaatt 480
112 gtggtgacag tcagggtgtc cctatatgac gaagataaca atggtgtaat ggatgaagg 540
113 aaggttatte catctgagac aatcgaggat gatatcaagg actgtgggct cttagacca 600
114 gatgttgaac tcgattatac gtggactcaa aacgagtgtg atctaccaga cacagtagac 660
115 gaggtgaag acacaccgtc agaaactgga gaattcttct ggtagatcta tcagaccact 720
116 tttatcagca ggacaactgg tcgttaccag acacctataa cgtgtcctca tcaataatgt 780
117 gtaaaacaga aataatcgat agaattattga aaataa 816
120 <210> SEQ ID NO: 3
121 <211> LENGTH: 852
122 <212> TYPE: DNA
123 <213> ORGANISM: Pholas dactylus
125 <400> SEQUENCE: 3
126 gtcggaaaag aacaaaatgg cttgtatcgt tttcgttgc cttgtcgtc tatgcttaat 60

```

RAW SEQUENCE LISTING

DATE: 03/05/2002

PATENT APPLICATION: US/09/831,142B

TIME: 14:22:53

Input Set : A:\Wcm-69us.app

Output Set: N:\CRF3\03052002\I831142B.raw

```

127 gcaaccgggt tccgggtgagg aagtacaatg cgcgatgaat tggacacaag ctaatgaata 120
128 tgtgttcaac gtggactgga tgaccatttt catctacgac tatggcgctc aagagcaact 180
129 gtacgaggat cgggcttttg ggctgtgtcg gattgaacgg gccggcccag gtaccacaaa 240
130 agccgtctgg attaactgga gtaacgacac gcagtcattgt gtaacaagaa aaacaatctt 300
131 cttcgagggt ggtggagaaa ttgcccggt agttgactac agaccacagg aagacggaac 360
132 tgagaaaact tttaacaagaa aattctctag caaaatgccg ggcaattaca tgcttatgga 420
133 cgtgtgcgct acaagggacg ctgatgataa atgcatcgaa ggcacaattg tggtgacagt 480
134 caggggtgtcc ctatatgacg aagataacaa tgggtgaatg gatgaaggta aggttattcc 540
135 atctgagaca atcgaggatg atatcaaggga ctgtgggctc ttagaccaag atgttgaaact 600
136 cgattatacg tggactcaaa acgagtgtga tctaccagac acagtagacg aggtcgaaga 660
137 cacaccgtca gaaactggag aattcttctg gtatgatctat cagaccactt ttatcagcag 720
138 gacaactggt cgttaccaga cacctataac gtgtcctcat caataatgtg taaaacagaa 780
139 ataatcgata gaatattgaa aataaaatgt taatagacac tggttgaaaa aaaaaaaaaa 840
140 aaaaaactcg ag 852
143 <210> SEQ ID NO: 4
144 <211> LENGTH: 225
145 <212> TYPE: PRT
146 <213> ORGANISM: Pholas dactylus
148 <400> SEQUENCE: 4
149 Met Ala Cys Ile Val Phe Val Ala Leu Val Ala Leu Cys Leu Met Gln
150 1 5 10 15
152 Pro Gly Ser Gly Glu Glu Val Gln Cys Ala Met Asn Trp Thr Gln Ala
153 20 25 30
155 Asn Glu Tyr Val Phe Asn Val Asp Trp Met Thr Ile Phe Ile Tyr Asp
156 35 40 45
158 Tyr Gly Ala Gln Glu Gln Leu Tyr Glu Asp Arg Ala Leu Gly Leu Cys
159 50 55 60
161 Arg Ile Glu Arg Ala Gly Pro Gly Thr Thr Lys Ala Val Trp Ile Asn
162 65 70 75 80
164 Trp Ser Asn Asp Thr Gln Ser Cys Val Thr Arg Lys Thr Ile Phe Phe
165 85 90 95
167 Glu Val Gly Gly Glu Ile Ala Arg Leu Val Asp Tyr Arg Pro Gln Glu
168 100 105 110
170 Asp Gly Thr Glu Lys Thr Phe Thr Arg Lys Phe Ser Ser Lys Met Pro
171 115 120 125
173 Gly Thr Tyr Met Leu Met Asp Val Cys Ala Thr Arg Asp Ala Asp Asp
174 130 135 140
176 Lys Cys Ile Glu Gly Thr Ile Val Val Thr Val Arg Val Ser Leu Tyr
177 145 150 155 160
179 Asp Glu Asp Asn Asn Gly Val Met Asp Glu Gly Lys Val Ile Pro Ser
180 165 170 175
182 Glu Thr Ile Glu Asp Asp Ile Lys Asp Cys Gly Leu Leu Asp Gln Asp
183 180 185 190
185 Val Glu Leu Asp Tyr Thr Trp Thr Gln Asn Glu Cys Asp Leu Pro Asp
186 195 200 205
188 Thr Val Asp Glu Ala Glu Asp Thr Pro Ser Glu Thr Gly Glu Phe Phe
189 210 215 220
191 Trp
192 225

```

RAW SEQUENCE LISTING

DATE: 03/05/2002

PATENT APPLICATION: US/09/831,142B

TIME: 14:22:53

Input Set : A:\Wcm-69us.app

Output Set: N:\CRF3\03052002\I831142B.raw

```

195 <210> SEQ ID NO: 5
196 <211> LENGTH: 205
197 <212> TYPE: PRT
198 <213> ORGANISM: Pholas dactylus
200 <400> SEQUENCE: 5
201 Glu Glu Val Gln Cys Ala Met Asn Trp Thr Gln Ala Asn Glu Tyr Val
202   1          5          10          15
204 Phe Asn Val Asp Trp Met Thr Ile Phe Ile Tyr Asp Tyr Gly Ala Gln
205          20          25          30
207 Glu Gln Leu Tyr Glu Asp Arg Ala Leu Gly Leu Cys Arg Ile Glu Arg
208          35          40          45
210 Ala Gly Pro Gly Thr Thr Lys Ala Val Trp Ile Asn Trp Ser Asn Asp
211          50          55          60
213 Thr Gln Ser Cys Val Thr Arg Lys Thr Ile Phe Phe Glu Val Gly Gly
214  65          70          75          80
216 Glu Ile Ala Arg Leu Val Asp Tyr Arg Pro Gln Glu Asp Gly Thr Glu
217          85          90          95
219 Lys Thr Phe Thr Arg Lys Phe Ser Ser Lys Met Pro Gly Thr Tyr Met
220          100         105         110
222 Leu Met Asp Val Cys Ala Thr Arg Asp Ala Asp Asp Lys Cys Ile Glu
223          115         120         125
225 Gly Thr Ile Val Val Thr Val Arg Val Ser Leu Tyr Asp Glu Asp Asn
226          130         135         140
228 Asn Gly Val Met Asp Glu Gly Lys Val Ile Pro Ser Glu Thr Ile Glu
229 145          150         155         160
231 Asp Asp Ile Lys Asp Cys Gly Leu Leu Asp Gln Asp Val Glu Leu Asp
232          165         170         175
234 Tyr Thr Trp Thr Gln Asn Glu Cys Asp Leu Pro Asp Thr Val Asp Glu
235          180         185         190
237 Ala Glu Asp Thr Pro Ser Glu Thr Gly Glu Phe Phe Trp
238          195         200         205
241 <210> SEQ ID NO: 6
242 <211> LENGTH: 225
243 <212> TYPE: PRT
244 <213> ORGANISM: Pholas dactylus
246 <400> SEQUENCE: 6
247 Met Ala Cys Ile Val Phe Val Ala Leu Val Ala Leu Cys Leu Met Gln
248   1          5          10          15
250 Pro Gly Ser Gly Glu Glu Val Gln Cys Ala Met Asn Trp Thr Gln Ala
251          20          25          30
253 Asn Glu Tyr Val Phe Asn Val Asp Trp Met Thr Ile Phe Ile Tyr Asp
254          35          40          45
256 Tyr Gly Ala Gln Glu Gln Leu Tyr Glu Asp Arg Ala Leu Gly Leu Cys
257          50          55          60
259 Arg Ile Glu Arg Ala Gly Pro Gly Thr Thr Lys Ala Val Trp Ile Asn
260  65          70          75          80
262 Trp Ser Asn Asp Thr Gln Ser Cys Val Thr Arg Lys Thr Ile Phe Phe
263          85          90          95
265 Glu Val Gly Gly Glu Ile Ala Arg Leu Val Asp Tyr Arg Pro Gln Glu

```

RAW SEQUENCE LISTING

DATE: 03/05/2002

PATENT APPLICATION: US/09/831,142B

TIME: 14:22:53

Input Set : A:\Wcm-69us.app

Output Set: N:\CRF3\03052002\I831142B.raw

```

266           100           105           110
268 Asp Gly Thr Glu Lys Thr Phe Thr Arg Lys Phe Ser Ser Lys Met Pro
269           115           120           125
271 Gly Thr Tyr Met Leu Met Asp Val Cys Ala Thr Arg Asp Ala Asp Asp
272           130           135           140
274 Lys Cys Ile Glu Gly Thr Ile Val Val Thr Val Arg Val Ser Leu Tyr
275 145           150           155           160
277 Asp Glu Asp Asn Asn Gly Val Met Asp Glu Gly Lys Val Ile Pro Ser
278           165           170           175
280 Glu Thr Ile Glu Asp Asp Ile Lys Asp Cys Gly Leu Leu Asp Gln Asp
281           180           185           190
283 Val Glu Leu Asp Tyr Thr Trp Thr Gln Asn Glu Cys Asp Leu Pro Asp
284           195           200           205
286 Thr Val Asp Glu Ala Glu Asp Thr Pro Ser Glu Thr Gly Glu Phe Phe
287           210           215           220

```

289 Trp

290 225

293 <210> SEQ ID NO: 7

294 <211> LENGTH: 17

295 <212> TYPE: DNA

296 <213> ORGANISM: Artificial sequence

298 <220> FEATURE:

299 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic

300 oligonucleotide

302 <220> FEATURE:

303 <221> NAME/KEY: modified_base

304 <222> LOCATION: (3)

305 <223> OTHER INFORMATION: i

307 <400> SEQUENCE: 7

W--> 308 acnathhttyt tycargt

17

311 <210> SEQ ID NO: 8

312 <211> LENGTH: 17

313 <212> TYPE: DNA

314 <213> ORGANISM: Artificial sequence

316 <220> FEATURE:

317 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic

318 oligonucleotide

320 <220> FEATURE:

321 <221> NAME/KEY: modified_base

322 <222> LOCATION: (12)

323 <223> OTHER INFORMATION: A, T, C or G

325 <220> FEATURE:

326 <221> NAME/KEY: modified_base

327 <222> LOCATION: (15)

328 <223> OTHER INFORMATION: i

330 <400> SEQUENCE: 8

W--> 331 cargargarg gnacnga

17

334 <210> SEQ ID NO: 9

335 <211> LENGTH: 17

Use of n and / or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to ensure a corresponding explanation is present in the <220> to <223> fields of each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/831,142B

DATE: 03/05/2002

TIME: 14:22:55

Input Set : A:\Wcm-69us.app

Output Set: N:\CRF3\03052002\I831142B.raw

L:308 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:331 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:354 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9
L:372 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10
L:557 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23